

Cherkasky – Fusionsproteine enthaltend Antikörperbinde -, Antigenbinde - Mikrotubulibinde und Immunantwortauslösende Regionen

Die Erfindung betrifft die Bereiche der Tumorphysiologie und der Biotechnologie.

In der Tumorthherapie stellen Operationen, Bestrahlung und Chemotherapie nach wie vor die entscheidenden Maßnahmen zur Therapie der Erkrankung dar. Bei der chemischen Tumorthherapie (Chemotherapie) werden je nach Tumortyp meist Zytostatika unterschiedlicher Wirkungsart verwendet, so etwa Alkylantien, Nitrosoharnstoffverbindungen, Folsäureantagonisten, Pyrimidin- und Purinanaloga wie Fluorouracil, Antibiotika mit Wirkung auf die DNA-abhängige RNA-Polymerase oder Enzyme wie L-Asparaginase. Eine Gruppe von Cytostatika für die Chemotherapie sind die Mitosehemmstoffe wie etwa Taxol und Vinca-Alkaloide.

Auf Grund ihrer sehr guten Antitumor-Aktivität haben besonders die Mitosehemmstoffe in letzter Zeit verstärkte Beachtung gefunden. Die Mitosehemmer beeinflussen den Aufbau oder Abbau der aus Mikrotubuli bestehenden Teilungsspindel – und greifen somit an der Zellteilung an. Das bekannte Colchicin oder Vinca-Alkaloide binden an spezifischen Bindestellen des α - oder β -Tubulins – als Baustein der Mikrotubuli – und bewirken z.B. eine Hemmung des Aufbaus der Mikrotubuli. Andere Mitosegifte – beispielsweise das Taxol – bewirkt deren Destabilisierung.

Mitose- oder Spindelgifte sind hochgradig toxisch und sind daher für therapeutische Zwecke problematisch. Die Toxizität von Colchicin ist sogar so hoch, daß diese Substanz bislang gar nicht therapeutisch verwendet wird. Das aus Eiben (*Taxus*) isolierte Alkaloid Taxol ist derzeit Gegenstand intensiver Forschung.

Die meisten Mitosehemmer binden an das β -Tubulin der Mikrotubuli. Dazu weisen sie Bindungsstellen auf, deren unterschiedliche hohe Spezifität für eine Klassifizierung der Mitosehemmer herangezogen wird. So werden verschiedene Gruppen wie der Colchizin-Typ, der Taxan-Typ, der Vinca Alkaloid Typ oder der Rhyoxin Typ unterschieden.

Auf Grund der hohen Toxizität der Zytostatika ist eine Therapie mit diesen Substanzen mit vielen Nebenwirkungen verbunden, die für die betroffenen Patienten oft kaum erträglich sind. Daher wird seit vielen Jahren an der Verbesserung der Therapien mit der Zielsetzung der Vermeidung oder Reduzierung der Nebenwirkungen gearbeitet. Ein Ansatz dazu stellt der Versuch dar, die Wirkstoffe gezielt nur zu den zu therapierenden Zellen – d.h. zu den Zielzellen – zu lenken.

Eine Möglichkeit zur Verwirklichung dieses Ansatzes basiert im wesentlichen darauf, zelltyp-spezifische Epitope zu identifizieren, einen Epitop-spezifischen monoklonalen Antikörper zu erzeugen und den derart gewonnen Antikörper oder Antigen-bindende Fragmente davon mit einem therapeutisch wirksamen Molekül zu koppeln. Ein derartiger Ansatz ist Gegenstand eines Forschungsprojekts der Universität von Kalifornien mit dem Ziel einer spezifischen Therapie von Brustkrebs (Sherie L. Morrison, Ph.D.: "Antibody Fusion Proteins for the

Therapy of Breast Cancer", University of California, Los Angeles, 1997-1999). Hierbei wurden Antikörper gegen die brustkrebspezifischen Moleküle her2/neu und CEA verwendet und mit immunstimulierenden Molekülen verbunden, welche die Aktivität der T-Zellen stimulieren.

Obwohl dieser Ansatz mit dem Vorteil einer hohen therapeutischen Selektivität einhergeht, ist er in der Praxis nur unter großen Anstrengungen bei hohem Aufwand und langer Entwicklungsdauer umzusetzen, da zahlreiche Entwicklungsschritte zu seiner Realisierung erforderlich sind. Hierzu müssen zunächst für den jeweiligen Zelltyp spezifische Antigene isoliert werden. Da es sich bei diesen in der Regel um Proteinantigene handelt, werden im folgenden zellspezifische Epitope des Antigens ermittelt, die möglichst geringe Ähnlichkeiten zu Epitopen der Proteine anderer Zelltypen aufweisen. Dies ist erforderlich zur Vermeidung von Kreuzreaktivitäten der therapeutisch eingesetzten Antikörper. Anschließend erfolgt die Herstellung monoklonaler, gegen das jeweilige Antigen gerichteter Antikörper, die im weiteren aufwendigen Selektions- und/oder Mutageneseverfahren wie etwa Phage Display unterzogen werden müssen, um zu einem Antikörper möglichst hoher Spezifität, bzw. möglichst geringer Kreuzreaktivität zu gelangen.

Darüber hinaus ergeben sich häufig Schwierigkeiten bei der Herstellung des gebrauchsfertigen Therapeutikums, da ein nicht-humaner Antikörper modifiziert werden muß, um ohne hohes allergenes Potential eingesetzt werden zu können. Dazu können die variablen Regionen, insbesondere jedoch die *Complementary determining regions* (CDR) in ein humanes Antikörpergerüst eingesetzt, wobei im fertigen Therapeutikum unterschiedlich große antigenspezifische Elemente des therapeutischen Antikörpers, so etwa die antigenbindenden Fragmente (Fab) zum Einsatz kommen. Dabei handelt es sich in aller Regel um antigenspezifische Elemente, die mindestens aus zwei separaten Polypeptidketten bestehen. Die Herstellung dieser komplexen antigenspezifischen Elemente und ihre Verknüpfung mit dem eigentlich therapeutischen Molekül ist in der Praxis oft aufwendig und erfordert komplexe Expressionskonstrukte und entsprechend geeignete Wirtszellen.

Bekannt sind wissenschaftliche Arbeiten, in welchen Fusionsproteinen bestehend aus Liganden und Mikrotubuli - Bindedomänen, sowie aus Liganden, Mikrotubuli - Bindedomänen und Membranpenetrationsdomänen beschrieben sind. (DE 199 25 052.9; DE 101 61 899.9; DE 101 61 738.0; DE 101 61 739.9 und DE 101 62 870.6)
Die Nachteile bestehen darin dass erstens, keine zusätzliche Verstärkung der Wirkung durch Auslösung einer Immunantwort erfolgt und zweitens die Auswahl an Tumorspezifischen Antigenen und Antigenbinderegionen relativ gering ist und man benötigt eine Komplexierung mit Antikörper, um jeden beliebigen Antikörper modifizieren zu können, um ihn fähig zu machen, direkt nach der Penetration oder Internalisierung Cytoskelett - Bestandteile bzw. Mikrotubuli zu binden bzw. zu fesseln.

Die Aufgabe der Erfindung besteht darin, effektive und selektive neuartige Fusionsproteinen und Fusionsprotein - Antikörper - Komplexen gegen unterschiedliche Arten der Leukämien und solide Tumoren zu entwickeln.

Die Aufgabe der Erfindung wird durch Fusionsproteinen, Cherkasky - Fusionsproteine genannt, enthaltend Antikörperbinde -, Antigenbinde -, Mikrotubulibinde - und Immunantwortauslösende Regionen gelöst. Die Selektivität wird durch Zell - oder Tumor spezifischen Liganden der Fusionsproteinen enthaltend zusätzlich Mikrotubuli - Binde und Immunantwort auslösenden Regionen sowie durch Antikörper der Fusionsprotein - Antikörper - Komplexen erreicht.

Die Fusionsproteinen, die mit Antikörpern Komplexe bilden enthalten Antikörper – Binde – und Mikrotubulibinderegionen.

Die Wirkung dieser Fusionsprotein – Antikörper – Komplexen besteht darin, Mikrotubuli - bzw. Zytoskelett zu finden bzw. zu fesseln nach dem die Tumorzellen durch hochaffine Antikörper ihre Zielzellen binden und penetrieren.

Die Antikörper Binderegion ist z.B. Staphylokokken protein A (SPA), extrazelluläre Region des Fc Rezeptors CD 64 etc.

Die Mikrotubuli - Binderegion ist z.B. Gephyrin, putativen Mikrotubulibindeprotein, FLJ 31424 Fis, MID - 1, MAP, Tau etc.

Diese Fusionsproteine können außerdem lange und superlange Spacer bzw. - Linkerregionen wie z.B. Polyglyzin oder Polyprolin enthalten, die mit Membranpenetrationsdomäne (MBD) oder Proteintransduktionsdomäne (PTD) fusioniert sind.

Die Fusionsproteine können auch Immunantwort - auslösenden Regionen wie z.B. Fc, B 7.1 oder B 7.2 enthalten, um die Wirkung der Komplexe zu erhöhen.

Die Fusionsproteine können Nukleinsäure – wie z.B. RecA oder Polysaccharid - Bindedomänen wie z.B. die Cellulose binderegion des C1PA enthalten, bzw. mit denen fusioniert werden, um durch Verdichtung bzw. erhöhte Konzentration die Wirkung zu verstärken.

Die Fusionsproteine können mit GFP oder anderen fluoreszenten Proteinen fusioniert werden, um ihre Wirkung optisch zu verfolgen oder ihre Konzentration durch Intensität der Fluoreszenz zu messen.

Außerdem können diese Fusionsproteine Gelenkregionen, wie z.B. Fünf - Glyzinregionen und mindestens eine GST-, Histag oder eine andere Region zur Durchführung der Affinitätsreinigung enthalten.

Die Fusionsproteine enthaltend Antigenbinde -, Mikrotubulibinde – und Immunantwort auslösenden Regionen entfalten ebenfalls eine Doppelwirkung.

Die Wirkung dieser Fusionsproteine besteht in der Tumorzellspezifischen Internalisierung und Hemmung der Zellteilung durch Bindung der Mikrotubuli sowie in der Auslösung einer Tumorzellspezifischer Immunantwort. Die Mikrotubulibinderegion kann z.B. aus Gephyrin, putativen Mikrotubulibindeprotein, FLJ 31424 Fis, Tau, MID - 1 oder MAP 1 ausgewählt werden und die Immunantwort auslösende Region ist z. B. Fc - Region eines IgG - Antikörpers, B 7.1 oder B7.2 Regionen zur Auslösung einer T - Zell - Reaktion.

Die Antigenbinderegionen können vorzugsweise aus folgenden Proteinen - EGF, FGF, CSF, MGF, IL – 15, IL – 2, etc. ausgewählt werden.

Diese Fusionsproteinen können entweder eine Proteintransduktions (PTD) oder keine PTD enthalten. PTD ist dann nicht nötig, wenn die Antigenbinderegion ein Ligand darstellt welches nach der Interaktion mit dem entsprechenden Rezeptor in die Zelle internalisiert wird. In anderen Fällen dient PTD dazu Fusionskonstrukt in die Zielzellen zu bringen, bzw. zu internalisieren. Die eine PTD bzw. MPD ist z.B. die von Gen – 3 – Protein des Bakteriophagen fd, gp 41 oder Tat Protein HIV – 1.

Die beschriebene Proteintransduktionsdomäne (PTD) ist eine elf - Aminosäure lange Region, die eine Region des HIV Tat Proteins darstellt.

Dem Forscher Dowdy und seinen Kollegen ist gelungen, 60 Proteine in der Größen Ordnung zwischen 15 kDa und 120 kDa zu fusionieren und nach folgender Denaturierung der Fusion

mit Harnstoff ins Zytosol zu transportieren. (Science (285, 1569 - 1572, 1999) und Nature biotechnology Vol. 17 S. 942, Oct. 1999).

Nach der Internalisierung des Fusionsproteins wirkt die Mikrotubuli - Bindedomäne wie z.B. Gephyrin, Tau, MAP oder MTD - 1 im Zytosol. Sie bindet Mikrotubuli und fesselt somit das Zytoskelett. Das dynamische Gleichgewicht (Wilde et al Nature cell biology 2001, March, vol. 3 und Carazo. Salas et al. Nature Cell biology 2001, March, vol.3.) der Mikrotubuli wird beeinträchtigt und die jeweilige Zelle kann sich nicht mehr teilen. Sobald sie sich nicht mehr teilt, stirbt sie .

Dadurch wird das Wachstum des Tumors, z.B. eines soliden Tumors gehemmt.

Die Fusionsproteine können zusätzlich GFP oder eine andere fluoreszente Region enthalten um die Wirkung optisch zu verfolgen und die Konzentration in einer Lösung durch Intensität der Fluoreszenz zu messen. Ausserdem können diese Fusionsproteine Gelenkregionen, vorzugsweise Fünf - Glyzin - Region, und mindestens ein GST -, His tag oder eine andere Region zur Durchführung der Affinitätsreinigung enthalten.

In der Seq. 1a ist die Aminosäuresequenz des Fusionsproteins SPA – 5G -- Gephyrin dargestellt.

In der Seq. 1b ist die Nukleinsäure kodierend für dieses Fusionsprotein dargestellt.

In der Seq. 2a ist die Aminosäuresequenz des Fusionsproteins SPA – 5G – Mikrotubuli Bindeprotein (MBP) dargestellt.

In der Seq. 2b ist die Nukleinsäuresequenz kodierend für das Fusionsprotein in der Fig. 2a dargestellt.

In der Seq. 3 ist die Nukleinsäure kodierend für das Fusionsprotein SPA – 5G – FLJ 31424 fis dargestellt.

In der Seq. 4 ist die Nukleinsäuresequenz kodierend für das Fusionsprotein IL 15 – 5G – Gephyrin - Fc dargestellt.

In der Seq. 5 ist die Nukleinsäuresequenz kodierend für das Fusionsprotein IL 2 – 5G – Gephyrin - Fc dargestellt.

In der Seq. 6 ist die Nukleinsäuresequenz kodierend für das Fusionsprotein IL 15 – 5G – MBP - Fc dargestellt.

In der Seq. 7 ist die Nukleinsäuresequenz kodierend für das Fusionsprotein IL 2 – 5G – MBP - Fc dargestellt.

In der Seq. 8 ist die Nukleinsäuresequenz kodierend für das Fusionsprotein IL 15 – 5G – FLJ 31424 fis - Fc dargestellt.

In der Seq. 9 ist die Nukleinsäuresequenz kodierend für das Fusionsprotein IL 2 – 5G – FLJ 314424 fis - Fc dargestellt.

In der Seq. 10 ist die Nukleinsäuresequenz des Fusionsproteins SPA – 5G – MBP - Fc dargestellt.

In der Seq. 11 ist die Nukleinsäuresequenz des Fusionsproteins SPA – 5G – Gephyrin - Fc dargestellt.

In der Seq. 12 ist die Nukleinsäuresequenz des Fusionsproteins SPA – 5G – FLJ 314424 fis - Fc dargestellt.

Die Wirkung der Fusionsproteine in den Sequenzen 1a – 3 besteht darin, dass sie einen beliebigen therapeutisch wirksamen Antikörper binden und dadurch modifizieren können. Die durch Bindung der Fusionsproteinen an Antikörper entstandenen Fusionsprotein – Antikörper – Komplexe können nach der Bindung der entsprechenden Antigenen und nach nachfolgender Internalisierung oder Penetration in den Zellen Mikrotubuli binden somit fesseln und dadurch die Zellteilung der bösartig veränderten Zellen hemmen.

Die Wirkung der Fusionsproteinen in den Sequenzen 4 – 9, besteht in der zielgerichteten Bindung ihrer Antigene oder Zell – bzw. Tumor spezifischen Rezeptoren, wonach die Doppelwirkung der Mikrotubulibinderegionen und zusätzlich die Induktion einer Immunreaktion durch Makrophagen entfalten wird bzw. erfolgt.

Die Wirkung der Fusionsproteinen in den Sequenzen 10 – 12, besteht in der Modifikation der Antikörper: die Komplexe besitzen auch eine Doppelwirkung und zwar die Fähigkeit Mikrotubuli zu binden und zusätzlich die Immunreaktion durch Aktivierung der Makrophagen zu induzieren.

Die DNA kodierend für die Fc Region des humanen Immunglobulins G ist von Nakamura, S., Sakugi, I., Kitai, K und Ichikawa, Y (NCBI [http://www. Ncbi. Nlm.nih.gov/entrez...NCBI Sequence Viewer](http://www.Ncbi.Nlm.nih.gov/entrez...NCBI Sequence Viewer),) beschrieben.

Die humane mRNA für PMBP (putative microtubuli – binding protein) ist von Nadezhdina E.S., beschrieben und ebenfalls in NCBI Sequence Viewer zu finden.

Die mRNA kodierend für Homosapiens Gephyrin (GPH) ist von NCBI beschrieben.

Die mRNA für Homo sapiens Interleukin – 2 ist von Chikara, S.K. und Sharma G beschrieben und ist in NCBI Sequence viewer zu finden.

Die Homo Sapiens mRNA für Interleukin 15 ist von Sorel, M.A. und Jacques, Y beschrieben wird in menschlichen Keratinocyten exprimiert und ist in NCBI Sequence Viewer zu finden.

Die Homo Sapiens cDNA für FLJ31421 fis, welche ein Mikrotubulibindeprotein kodiert, ist von Ota et al und Tashiro et al beschrieben und ist in NCBI Sequence Viewer zu finden.

Die SPA – Gensequenz bzw. die DNA kodierend für Staphylokokkenprotein A ist von El – Sayed, A., Alber J., Laemmer et al beschrieben und ist in NCBI Sequence Viewer zu finden.

Die Nukleinsäuren sind erhältlich von den Autoren. Alternativ dazu können die Sequenzen mit Hilfe von PCR und RT – PCR mit entsprechenden Primern herausamplifiziert und kloniert werden.

Diese Techniken sind dem Fachman bekannt.

Die Klonierung und Expression der Konstrukte wird vorzugsweise in E.Coli durchgeführt. Die frische E. Coli Zellkultur wird unter Zugabe von 75mM CaCl₂ (steril, kalt 250 ml) und Glycerin (steril, kalt 5mM) vorbereitet. Bei der Durchführung wird das Medium zuerst auf 37°C erwärmt, mit 8ml frischer E.Coli – Kultur angeimpft und bei 37°C kräftig geschüttelt. Danach wird die Kultur abgekühlt. Die Zellen werden 10 Minuten lang bei 6000 rpm im GSA – Rotor der Sorvall – Zentrifuge abzentrifugiert und in 200 ml eiskaltem CaCl₂ 875 mM)

suspendiert. Danach werden die Zellen 20min auf Eis gestellt und wieder 10 min bei 6000 rpm in GSA – Rotor (Sorwall – Zentrifuge) abzentrifugiert.

Die Zellen werden in 20 ml eiskaltem CaCl_2 (75 mM) resuspendiert und wieder für 15 min auf Eis gestellt. Dazu wird 4,2 ml Glycerin zugegeben und gemischt. Die Lösung wird in sterile Eppendorfgefäße zu 0,5 ml abgefüllt und bei -70°C eingefroren.

Die Transformation von E.Coli erfolgt mit 20 Mikroliter Ligationsansatz (oder max. 0,5 Mikroliter einer Plasmidpräparation, die 10 fach verdünnt wird). 100 ml kompetenter E.Coli Zellen werden zugegeben.

Danach werden diese ca. 30 min auf Eis (bis ca 1h) aufbewahrt, danach 2 – 3 min. auf $42 - 43^\circ\text{C}$ erwärmt, um den Hitzeschock hervorzurufen und dann wieder auf Eis gestellt. Die Zellen werden z.B. auf X – Gal, Lbamp ausplattiert.

Die Zell – PCR erfolgt mit folgendem PCR – Mix pro Ansatz: 9 Mikroliter H_2O , 10ml 10xPCR – Puffer 10 Mikroliter dNTP – Mix (2 bis 2,5 mM jedes), 5ml BSA (20 mg/ml) oder 5 Mikroliter H_2O zusätzlich, 2ml 5' – Primer (20 pmol / Mikroliter); 2ml 3, - Primer (20 pmol / Mikroliter) und 2 ml Taq – Pol (54 / Mikroliter); insgesamt 40 Mikroliter.

Eine kleine Menge Zellen wird von der Platte abgenommen, und in einem 1,5 ml Eppendorf – Gefäß unten an der Wand verteilt.

Danach werden die Proben 2 min in 600 W Mikrowelle mit offenen Eppendorf – Gefäßen hingelegt.

Dann wird 200 Mikroliter H_2O zugegeben und gut gevortext, um zu resuspendieren. Danach wird 1 min in Eppendorf – Zentrifuge zentrifugiert. Vom Überstand werden 60 Mikroliter in ein PCR – Eppendorfgefäß gegeben und 40 Mikroliter PCR mix zugegeben.

Die PCR wird gestartet. Die Zyklen sind wie üblich, je nach annealing – Temperatur der Primer und Länge des zu Synthetisieren den DNA – Fragmentes, z. B.

94°C	2 min	45°C	1 min	72°C	2 min	1x
94°C	30 sec	45°C	30sec	72°C	2 min	4x
94°C	30 sec	58°C	30sec	72°C	2 min	32x
94°C	30 sec	58°C	30sec	72°C	5 min	1x

Die Ligation wird mit 20 Mikroliter - Ansätzen gemacht : 15 Mikroliter zu ligierende DNA in H_2O , 4 Mikroliter 5x Ligase – Puffer mit PEG 1 Mikroliter T4 DNA – Ligase (1U/ Mikroliter), insgesamt 20 Mikroliter. Sticky – end – Ligierungen 1 – 2,5 h bei Raumtemperatur und blunt – end – ligierungen 4h bei Raumtemperatur. Ligase – Puffer besteht aus 250 mM tris/HCL pH 7, 6, 50 mM MgCl_2 , 25% PEG 6000 (Sigma oder Serva) oder PEG 8000, 5mM ATP und 5mM DTT.

Die E. coli wird also transformiert.

Ausführungs - Beispiel 1

Klonierung und Expression des Fusionskonstruktes
N - IL - 15 - L - Gephyrin - Fc - C

c DNA für Gephyrin wird durch PCR kloniert, bzw. die Homosapiens Gephyrin (GPH) mRNA wird durch RT - PCR kloniert. Die Daten der GPH mRNA - Sequenz sind beim National Center für Biotechnology Information, NIH, Bethesda MD, 20 894, USA erhältlich, sowie auf der Internet - Seite von NCBI ([http:// www. ncbi. Nlm.nih.go v](http://www.ncbi.nlm.nih.gov)) zu finden.

Die IL - 15 mRNA sowie die Fc vom IgG - mRNA werden ebenfalls durch RT - PCR kloniert.

Das Fusionsprotein bzw. das Fusionsprodukt wird aus PCR - Produkten zusammen gesetzt. Die PCR - Primern sind so konstruiert, dass sie Restriktionsstellen auf 5' und 3' Enden enthalten, um spätere Ligationsschritte durchzuführen. Die 5' und 3' Endes des IL - 15 PCR

Produkts enthalten Bam HI und Hind III Restriktionsstellen. Die 5' und 3' Enden des Gephyrin - PCR - Produktes enthalten EcoR I und KpnI Restriktionsstellen und die 5' und 3' Endes des Fc - PCR - Produktes enthalten Pst I und Sac I Restriktionsstellen.

Ligation der IL - 15, Gephyrin und Fc - sequenzen in den pUC 19 (2686 bp) - Vektor erfolgt unter Standard - bedingungen. Der PUC 19 - Vektor wird zuerst mit Bam HI und Hind 3 behandelt.

Der IL - 15 Segment wird durch diese Behandlung in den Vektor hineinligiert, pUC 19 - IL - 15 wird mit Eco RI und Kpn I behandelt, um den Gephyrin - Segment hineinzuligieren. Der pUC 19 - IL - 15 - Gephyrin - Vektor wird abschliessend mit Pst I und Sac I behandelt, um Fc Segment in den Vektor hineinzuligieren.

Ligationspuffer wird aus 66 mM Tris, pH 7,6, 5 mM DTT und 1mM ATP sowie aus 20 Mikroliter T4 - DNA Ligase zusammengesetzt.

Das Ligationprodukt wird in E. Coli transformiert, exprimiert und abschliessend gereinigt.

Ausführungs - Beispiel 2

Klonierung und Expression des Fusionskonstruktes SPA - 5G - Gephyrin

C DNA für Gephyrin wird durch PCR kloniert, bzw. die Homo sapiens Gephyrin (GPH) mRNA wird durch RT - PCR kloniert. Die Daten der GPH mRNA - Sequenz sind bei im Internet, beim National Center for Biotechnology information, NIH, Bethesda MD 208 94, USA erhältlich, sowie auf der Internet Seite von NCBI ([http:// www. ncbi.nlm. nih.gov: 80/entez/... eotide...](http://www.ncbi.nlm.nih.gov:80/entrez/... eotide...)) zu finden.

C DNA für SPA ligiert mit dem Primer kodierend für die Fünf - Glyzin - Spacer wird ebenfalls mit PCR kloniert.

Das Fusionsprotein wird aus PCR - Produkten zusammengesetzt. Die PCR - Primern sind so konstruiert, dass sie Restriktionsstellen auf 5' und 3' Enden enthalten, um spätere Ligationsschritte durchzuführen. Die 5' und 3' Enden des Gephyrin - PCR - Produkts enthalten Bam HI und Hind III Restriktionsstellen. Die 5' und 3' Enden des SPA - PCR - Produkts enthalten XmnI und Bg III Restriktionsstellen.

Nach der Amplifikation und Reinigung werden die PCR Produkten in PCR II Vektoren ligiert. Positive Klone werden durch Screening Plasmide richtiger Masse identifiziert. Die Klone werden durch DNA - Sequenzierung bzw. durch Standardmethoden überprüft bzw. bestätigt.

Das Gephyrin - PCR - Produkt wird aus der PCR II durch restriktive Spaltung durch Bam HI und Hind III herausgeschnitten, und SPA - PCR - Produkt wird aus dem PCR II durch Xmn I und Bg III herausgeschnitten.

Ligation der Gephyrin und SPA - Segmente in den pMal - c 2 Expressionsvektor erfolgt unter Standard - Bedingungen. Der p Mal - c 2 Vektor wird mit Bam H I und Hind 3 behandelt. Der Gephyrin - Segment wird durch diese Behandlung in den pMal - c 2 hinein ligiert.

PMal - c 2 - Gephyrin wird mit X mnI und Bam HI geschnitten, um SPA Segment hinein zu ligieren.

Ligationspuffer wird aus 66 mM Tris PH 7,6, 5 mM Mg Cl₂, 5 mM DTT und 1 mM ATP, sowie aus der T4 DNA Ligase (insgesamt 20 Mikroliter) zusammengesetzt. Die Ligation wird bei 14°C durchgeführt.

Das Ligationsprodukt wird in E Coli transformiert, exprimiert und abschliessend gereinigt.

Mit den in den Ausführungsbeispielen 1 und 2 beschriebenen Methoden sowie mit den anderen dem Fachman bekannten Methoden werden die Sequenzen in den Figuren 1 – 12 kloniert und exprimiert.

Cherkasky – Fusionsproteine enthaltend Antikörperbinde -, Antigenbinde - Mikrotubulibinde und Immunantwortauslösende Regionen

Patentansprüche

Beansprucht werden:

1. Fusionsproteine, dadurch gekennzeichnet, dass sie spezifische Antigenbinderegionen wie vorzugsweise EGF, FGF, CSF, MGF, IL – 15, IL – 2, oder anderen Liganden oder ihre Regionen, Mikrotubulibinderegionen wie vorzugsweise Gephyrin, Tau, MAP, MID – 1, MBP, bzw. put MBP oder PMBP, FLJ 31424 Fis, oder ihre Regionen und Immunantwortauslösenden Regionen wie vorzugsweise Fc des IgG, B 7.1, B 7.2 oder ihre Regionen, enthalten.
2. Fusionsproteine, dadurch gekennzeichnet, dass sie Antikörperbinderegionen, wie vorzugsweise Staphylokokkenprotein A (SPA), extra zelluläre Region des Fc Rezeptors CD 64 oder ihre Regionen und Mikrotubulibinderegionen, vorzugsweise Gephyrin, Tau, MAP, MID – 1 MBP, FLJ 31424 Fis, oder ihre Regionen enthalten.
3. Fusionsproteine nach dem Anspruch 2 gekennzeichnet durch immunauslösende Regionen wie vorzugsweise Fc Regionen des Immunglobulins G, oder IgG, HLA – B 7.1 oder HLA – B 7.2.
4. Fusionsproteine nach den Ansprüchen 1 – 3, gekennzeichnet durch lange und sehr lange Spacer oder Linker Regionen wie vorzugsweise Polyglycin, Polyprolin oder Spacer die Glycine und Proline enthalten.
5. Fusionsproteine nach den Ansprüchen 1 – 4, gekennzeichnet durch mindestens eine Nukleinsäurebinderegion vorzugsweise RecA.
6. Fusionsproteine nach den Ansprüchen 1 – 5, gekennzeichnet durch mindestens eine Polysaccharidbinderegion wie vorzugsweise die Cellulosebinderegion des CipA Proteins.
7. Fusionsproteine nach den Ansprüchen 1 – 6, dadurch gekennzeichnet dass sie GFP, eine andere fluoreszente Region, Membranpenetrationsdomäne wie vorzugsweise das Gen – 3 – Protein des Bakteriophagen fd, gp 41 oder Tat Protein des HIV – 1 oder eine mindestens andere Region, enthalten bzw. umfassen.
8. Fusionsproteine nach den Ansprüchen 1 – 7, gekennzeichnet durch eine GST Region, His tag oder eine andere Region zur Durchführung der Affinitätsreinigung.

9. Fusionsproteine nach den Ansprüchen 1 – 8, gekennzeichnet durch eine Spaltstelle für eine Protease vorzugsweise zum Abschneiden der GST oder eines anderen Reinigungstags vom Konstrukt.
10. Nukleinsäure und Aminosäure Sequenzen, DNA – Vektoren, Klonierungs – und Expressionssysteme für die Fusionsproteine nach den Ansprüchen 1 – 9, sowie alle Aminosäure und Nukleinsäuresequenzen oder Sequenzen 1 – 12 .

Seq. 1a

AAQHDEAQQNAFYQVLNMPNLNADQRNGFIQSLKDDPSQSANVL
GEAKKLNESQAPKADNNFNKEQQNAFYEILNMPNLNEEQRNGFIQSLKDDPSQSANLL
SEAKKLNESQAPKADNKFNKEQQNAFYEILHLPNLNEEQRNGFIQSLKDDPSQSANLL
AEAKKLNDQAAPKADNKFNKEQQNAFYEILHLPNLTEEQRNGFIQSLKDDPSVSKEIL
AEAKKLNDQAAPKEEDNNKPGKEDGNKPGKEDGNGGGGG
MSPFPLTSMDKAFITVLEMTPVLGTEI:NYRDGMGRVLAQDVYA
KDNLPPFPASVKDGYAVRAADGPGDRFIIGESQAGEQPTQTVMPGQVMRVTTGAPIPC
GADAVVQVEDTELIRESDDGTEELEVRILVQARPGQDIRPIGHDIKRGEVLAKGTHM
GPSEIGLLATVGVTEVEVNKFPVVAVMSTGNELLPEDDLLPGKIRDSNRSTLLATIQ
EHGYPTINLGIVGDNPDLLNALNEGISRADVITSGGVSMGEKDYLKQVLDIDLHAQ
IHFGRVFMKPGLPTTFATLDIDGVRKIIFALPGNPVSAVVTCLFVVPALRKMQGILD
PRPTIIKARLSCDVKLDPRPEYHRCILTWHHQEPLPWAQSTGNQMSSRLMSMRSANGL
LMLPPKTEQYVELHKGEVVDVMVIGRL

Seq. 1b

tgctgcgcaa cacgatgaag ctcaacaaaa cgctttttat caagtcttaa atatgcctaa
cttaaatgct gatcaacgca atggttttat ccaaagcctt aaagatgac caagccaaag
tgctaacgtt ttaggtgaag ctaaaaaatt aaacgaatct caagcaccga aagctgacaa
caatttcaac aaagaacaac aaaatgcttt ctatgaaatc ttgaacatgc ctaactgaa
cgaagaacaa cgcaatgggt tcatccaaag cttaaaagat gacccaagtc aaagtgcctaa
cctattgtca gaagctaaaa agttaaatga atctcaagca ccgaaagcgg ataacaaatt
caacaaagaa caacaaaatg ctttctatga aatcttacat ttacctaact taaacgaaga
acaacgcaat gggttcatcc aaagcctaaa agatgaccca agccaaagcg ctaacctttt
agcagaagct aaaaagctaa atgatgcaca agcaccaaaa gctgacaaca aattcaacaa
agaacaacaa aatgctttct atgaaatctt acatttacct aacttaactg aagagcaacg
taacggcttc atccaaagcc ttaaagacga tccttcagt agcaaagaaa ttttagcaga
agctaaaaag ctaaacgatg ctcaagcacc aaaagaggaa gacaacaaca aacctggtaa
agaagacggc acaaacctg gcaaagaaga cggtaacggc ggc ggc ggc ggc
gtttaggtca cagtgcgtc gatcacca aggtggctag aagacatgc atgtctcctt
ttcctctgac atctatggac aaagccttta tcacagtcct ggagatgact ccggtgctt
ggacagaaat catcaattac cgagatggaa tggggcgagt ccttgctcaa gatgtatatg
caaaagacaa ttaccccc ttccagcat cagtaaaaga tggctatgct gtccgagctg
ctgatggccc aggagatcgt tcatcattg gggaatcca agctggtgaa cagccaactc
agacagtaat gccaggacaa gtcatgcggg ttacaacagg tgctccaata cctgcggtg
ctgatgcagt agtacaagt gaagataccg aacttatcag ggaatcagat gatggcactg

aagaacttga agtgcaatt ctggtgcaag ctgggccagg ccaagatatc agacccatcg
 gccatgacat taaaagaggg gaatgtgtt tggccaaagg aaccacatg ggcccctcag
 agattggtct tctggcaact gtaggtgtca cagagggtga agttaataag ttccagtgg
 ttgcagtcat gtcaacaggg aatgagctgc taaatcctga agatgacctc ttaccaggga
 agattcgaga cagcaatcgt tcaactcttc tagcaacaat tcaggaacat ggttacccca
 cgatcaactt gggatttga ggagacaacc cagatgactt actcaatgcc ttgaatgagg
 gtatcagtcg tgctgatgc atcatcacat cagggggtgt atccatgggg gaaaaggact
 atctcaagca ggtgctggac attgatcttc atgctcagat ccattttggc agggtttta
 tgaaaccagg cttgccaaca acatttgcaa ctttgatata tgatggtga agaaaaataa
 tctttgcaact acctgggaat cctgtatcgg ctgtggtcac ctgcaatctc ttgttgtgc
 ctgcactgag gaaaatgcag ggcacttgg atcctcggcc aacatcatc aaagcaaggt
 tatcatgtga tgtaaaactt gatcctcgtc cagaatacca tcggtgtata ctaactggc
 ataccaaga accactacct tgggcacaga gtacaggtaa tcaaatgagc agccgtctga
 tgagcatgcg cagtccaat ggattgtga tgctacctc aaagacagaa cagtacgtg
 agctccacaa aggcgaggtg gtggatgtca tggtcattgg acggctatga tggtcaccag

Seq. 2a

AAQHDEAQQNAFYQVLNMPNLNADQRNGFIQSLKDDPSQSANVL
 GEAKKLNESQAPKADNNFNKEQQNAFYEILNMPNLNEEQRNGFIQSLKDDPSQSANLL
 SEAKKLNESQAPKADNKFNKEQQNAFYEILHLPNLNEEQRNGFIQSLKDDPSQSANLL
 AEAKKLNDAPKADNKFNKEQQNAFYEILHLPNLTEEQRNGFIQSLKDDPSVSKEIL
 AEAKKLNDAPKEEDNNKPGKEDGNKPGKEDGNGGGGG
 AAAXASTAKETA EAVADXLXKAGPLVAVSAVALDITAYP

Seq. 2b

tgctgcgcaa cacgatgaag ctcaacaaaa cgctttttat caagtcttaa atatgcctaa
 cttaaagtct gatcaacgca atggttttat ccaaagcctt aaagatgac caagccaaag
 tgctaactgt ttaggtgaag ctaaaaaatt aaacgaatct caagcaccga aagctgacaa
 caatttcaac aaagaacaac aaaatgcttt ctatgaaatc ttgaacatgc ctaactgaa
 cgaagaacaa cgcaatggtt tcatccaaag cttaaagat gaccaagtc aaagtgctaa
 cctattgtca gaagctaaaa agttaaatga atctcaagca ccgaaagcgg ataacaatt
 caacaaagaa caacaaaatg cttctatga aatcttacat ttacttaact taaacgaaga
 acaacgcaat ggtttcatcc aaagcctaaa agatgacca agccaaagcg ctaaccttt
 agcagaagct aaaaagctaa atgatgcaca agcaccacaaa gctgacaaca aattcaacaa
 agaacaacaa aatgctttct atgaaattt acatttacct aacttaactg aagagcaacg
 taacggcttc atccaaagcc ttaaagacga tccttcagt agcaaagaaa ttttagcaga

agctaaaaag ctaaacgatg ctcaagcacc aaaagaggaa gacaacaaca aacctggtaa
agaagacggc acaaacctg gcaaagaaga cggtaacggc ggc ggc ggc ggc
gcggccgct cgaccgcn gcgctcgacg gcaaaggaga ctgctgaggc tgttgctgat
ganatactgg anaaggctgg gccactgtt gctgtgtctg ctgtgcact tgatataact
gcctaccct aaaagccaaa

Seq. 3

tgctgcgcaa cacgatgaag ctcaacaaaa cgcttttat caagtcttaa atatgcctaa
cttaaatgct gatcaacgca atggttttat ccaaagcctt aaagatgac caagccaaag
tgctaacgtt ttaggtgaag ctaaaaaatt aaacgaatct caagcaccga aagctgacaa
caattcaac aaagaacaac aaaatgctt ctatgaaatc ttgaacatgc ctaactgaa
cgaagaacaa cgcaatggtt tcatccaaag cttaaaagat gacccaagtc aaagtgctaa
cctattgtca gaagctaaaa agttaaatga atctcaagca ccgaaagcgg ataacaatt
caacaaagaa caacaaaatg ctttctatga aatcttacat ttacctact taaacgaaga
acaacgcaat ggttcatcc aaagcctaaa agatgacca agccaaagcg ctacacctt
agcagaagct aaaaagctaa atgatgcaca agcaccaaaa gctgacaaca aattcaacaa
agaacaacaa aatgcttct atgaaattt acattacct aacttaactg aagagcaacg
taacggctc atccaaagcc ttaaagacga tccttcagt agcaaagaaa ttttagcaga
agctaaaaag ctaaacgatg ctcaagcacc aaaagaggaa gacaacaaca aacctggtaa
agaagacggc acaaacctg gcaaagaaga cggtaacggcggcgggcgggcggc
aatgtcccga attccagcc tcaccacccc ttctcagtaa tgacctgtt tgggtgcagg
aggtacctac tccatactga gggtgaaatt aagggaaggc aaagtccagg cacaagagt
ggacccagc ctctactct cagttccact catccaactg ggacctcac cacgaatct
atgatctgat tcggttccct gtctcctct cccgtcacag atgtgagcca gggcactgct
cagctgtgac cctaggtgt tctgcctgt tgacatggag agagccctt cccctgagaa
ggcctggccc ctctctgtc tgagcccaca gcagcaggct ggggtgtctt gttgtcagt
gtggcaccag gatggaagg caaggcacc agggcaggcc cacagtcccg ctgtcccca
cttgaccct agctgtagc tgccaacct ccagacagcc cagcccgctg ctacgtcca
catgcatagt atcagccct cacacccgac aaaggggaac acacccctt ggaaatggt
ctttccccc agtccagct ggaagccatg ctgtctgtc tgctggagca gctgaacata
tacatagatg ttgccctgcc ctccccatct gcacctgtt gagttgagt tggattgtc
tgttatgct tggattcacc agagtacta tgatagtga aagaaaaaaa aaaaaaaa
aggacgatg tatctgaaa tgctgtaaa gaggttcta acccaccct acgaggtgtc
tctaccccc acactgggac tctgtggcc tgtgtgtgc caccctgtg gggcctcca
agtttgaaa ggcttctc agcacctggg acccaacaga gaccagctc tagcagctaa
ggaggccgt cagctgtgac gaaggcctga agcacaggat taggactgaa gcgatgatg
cccttccct acttccctt ggggtccct gtgtcagggc acgactagg tctgtggct
ggctggctt gcggcgcgag gatggtctc tctggcata gccgaagtc tcatggcagt
cccaaaggag gcttacaact cctgcatcac aagaaaaagg aagccactgc cagctggggg

gatctgcagc tcccagaagc tccgtgagcc tcagccaccc ctacagactgg gttcctctcc
aagctcgccc tctggagggg cagcgcagcc tcccaccaag ggccctgcga ccacagcagg
gattgggatg aattgcctgt cctggatctg ctctagaggc ccaagctgcc tgcctgagga
aggatgactt gacaagtcag gagacactgt tcccaaagcc tggaccagag cacctcagcc
cgctgacctt gcacaaactc catctgctgc catgagaaaa gggaagccgc ctttgcaaaa
cattgctgcc taaagaaact cagcagcctc aggcccaatt ctgccacttc tggtttgggt
acagttaaag gcaaccctga gggacttggc agtagaaatc cagggcctcc cctggggctg
gcagcttcgt gtgcagctag agctttacct gaaaggaagt ctctgggccc agaactctcc
accaagagcc tccctgccgt tcgctgagtc ccagcaattc tctaagttg aagggatctg
agaaggagaa ggaaatgtgg ggtagatttg gtggtggta gagatatgcc cccctcatta
ctgccaacag ttctggctgc attcttcac gcacctcggg tctcttct gaagtcttg
tgccctgctc ttcagcacca tgggccttct tatacgaag gctctgggat ctcccccttg
tggggcaggc tcttggggcc agcctaagat catggttag ggtgatcagt gctggcagat
aaattgaaaa ggcacgctgg ctgtgatct taaatgagga caatcccccc agggctgggc
actctcccc tccctcact tctccacct gcagagccag tgccttggg tgggctagat
aggatatact gtatgccggc tcttcaagc tgcctgactc ctttatcaat agttccattt
aaattgactt cagtggtag actgtatct gtttgcatt gctgtgtg ctatggggg
aggggggagg aatgtgaag atagttaaca tgggcaaagg gagatcttg ggtgcagcac
ttaaactgcc tcgtaacct ttcatgatt tcaaccacat ttgctagagg gagggagcag
ccacggagtt agaggccctt ggggttctc tttccactg acaggcttc ccaggcagct
ggctagtca tccctcccc agccaggtgc aggcgtagga atatggacat ctggttgctt
tggcctgctg ccctcttca ggggtcctaa gccacaaac atgcctccct aagaccttg
catcttccc tctaagccgt tggcacctct gtgccacctc tcacactggc tccagacaca
cagcctgtgc ttttgagct gagatcactc gcttaccct cctcatctt gttctccaag
taaagccagc aggtcggggc gagggcagag gtgatccct gcgtgtccc tctacagacc
tgcagcttca taaaacttct gatttctct cagcttgaa aagggttacc ctgggcactg
gcctagagcc tcacctcta atagacttag ccccatgagt ttgcatgtt gagcaggact
atttctggca ctgcaagtc ccatgattc ttcggttaatt ctgaggggtg ggggagggac
atgaaatcat cttagcttag cttctgtct gtgaatgtct atatagtga ttgtgtgtt
taacaaatga ttacactga ctgtgtctg aaaagtgaat ttggaataa agttattact
ctcat

Seq. 4

ccatccagtg ctacttgtgt ttacttctaa acagtcattt tctaactgaa gctggcattc
atgcttcat tttgggctgt ttcagtgcag ggcttctaa aacagaagcc aactgggtga
atgtaataag tgatttgaaa aaaattgaag atctattca atctatgcat attgatgcta
cttatatac ggaaagtgat gttcaccoca gttgcaaagt aacagcaatg aagtgccttc
tcttgaggtt acaagttatt tcaattgagt ccggagatgc aagtattcat gatacagtag

aaaatctgat catcctagca aacaacagtt tgtcttctaa tgggaatgta acagaatctg
gatgcaaaga atgtgaggaa ctagaggaaa aaaatattaa agaattttg cagagtttg
tacatattgt ccaaagtgc atcaacactt ctt ggc ggc ggc ggc ggc
gtttaggta cagtgcgtc gatacacca aggtggctag aagacatgc atgtctctt
ttctctgac atctatggac aaagcctta tcacagtcct ggagatgact ccggtgctg
ggacagaaat catcaattac cgagatggaa tggggcgagt ccttgctcaa gatgtatatg
caaaagacaa ttaccccc ttcccagcat caglaaaaga tggctatgct gtccgagctg
ctgatggccc aggagatcgt tcatcattg gggaatccca agctggtgaa cagccaactc
agacagtaat gccaggacaa gtcatgcggg ttacaacagg tgctccaata ccctgcggtg
ctgatgcagt agtacaagtg gaagataccg aacttatcag ggaatcagat gatggcactg
aagaactga agtgcaatt ctggtgcaag ctgggccagg ccaagatctc agacccatcg
gccatgacat taaaagaggg gaatgtgtt tggccaaagg aaccacatg ggcccctcag
agattggtct tctggcaact gtaggtgtca cagaggttga agttaataag ttccagtgg
ttgcagtcat gtcaacaggg aatgagctgc taaatcctga agatgacctc ttaccaggga
agattcgaga cagcaatcgt tcaactctc tagcaacaat tcaggaacat ggttacccca
cgatcaactt gggatttga ggagacaacc cagatgactt actcaatgcc ttgaatgagg
gtatcagtcg tgctgatgc atcatcacat cagggggtgt atccatgggg gaaaaggact
atctcaagca ggtgctggac atgactctc atgtcagat ccattttggc agggtttta
tgaaaccagg ctgccaaca acatttgcaa ctttgatat tgatggtgta agaaaaataa
tcttgcaact acctgggaat cctgtatcgg ctgtggtcac ctgcaatctc ttgttgtgc
ctgcactgag gaaaatgcag ggcatttg atctcggcc aaccatcctc aaagcaagg
tatcatgtga tgtaaaact gatcctcgtc cagaatacca tcggtgtata ctaactggc
atcaccaaga accactacct tgggcacaga gtacaggtaa tcaaatgagc agccgtctga
tgagcatcg cagtccaat ggattgtga tgctacctc aaagacagaa cagtacgtg
agctccacaa aggcgagggtg gtggatgtca tggctattg acggctatga tggtcaccag
ctgttgacaa ttaatcatc gctcgtataa tgtgtggaat tgtgagcgga taacaattc
acacaggaaa caggatccga taatgacatg cccaccgtgc ccagcacctg aactcctggg
gggacgtca gtcttctct tcccccaaa acccaaggac acctcatga tctccggac
ccctgaggtc acatgcgtg tggtagcgt gagccacgaa gacctgagg tcaagttcaa
ctggtacgtg gacggcgtg aggtgcataa tgccaagaca aagccgcggg aggagcagta
caacagcacg taccgggtg tcagcgtct caccgtctg caccaggact ggctgaatg
caaggagtac aagtgaagg tctccaaca agccctcca gcccctatc agaaaacct
ctcaaagcc aaagggcagc cccgagaacc acaggtgtac acctgccc catcccggga
ggagatgacc aagaaccagg tcagcctgac ctgctggtc aaaggctct atccagcga
catgcccgtg gagtgggaga gcaatgggca gccggagaac aactacaaga ccacgcctcc
cgtgctggac tccgacggct ccttctct ctatagcaag ctaccgtgg acaagagcag
gtggcagcag gggaacgtct tctcatgctc cgtgatcat gaggctctg acaaccacta
cacgcagaag agcctctccc tgtcccggg taaataatag gatcc

Seq. 5

atgcctactt caagttctac aaagaaaaca cagctacaac tggagcattt actgctggat
ttacagatga tttgaatgg aattaataat tacaagaatc ccaaactcac caggatgctc
acatttaagt ttacatgcc caagaaggcc acagaactga aacatcttca gtgtctagaa
gaagaactca aacctctgga ggaagtgcta aatttagctc aaagcaaaaa ctttactta
agaccaggg acttaatcag caatatcaac gtaatagttc tggaactaaa gggatctgaa
acaacattca tgttgaata tgctgatgag acagcaacca tttagaatt totgaacaga
tggattacct tttctcaaag catcatctca aactgactt gataa ggc ggc ggc ggc ggc
gttaggtca cagtgtgtc gatatcacca aggtggctag aagacatgc atgtctcctt
ttctctgac atctatggac aaagccttta tcacagtcct ggagatgact ccggtgcttg
ggacagaaat catcaattac cgagatggaa tggggcgagt ccttgctcaa gatgtatatg
caaaagacaa ttaccccc ttccagcat cagtaaaaga tggctatgct gtccgagctg
ctgatggccc aggagatcgt tcatcattg gggaatcca agctggtgaa cagccaactc
agacagtaat gccaggacaa gtcatgcggg ttacaacagg tgctccaata ccctgcggtg
ctgatgcagt agtacaagt gaagataccg aactatcag ggaatcagat gatggcactg
aagaactga agtgcgaatt ctggtgcaag ctgggccagg ccaagatct agacccatcg
gccatgacat taaaagaggg gaatgtgtt tggccaaagg aaccacatg ggcccctcag
agattggtct totggcaact gtaggtgtca cagaggttga agttaataag ttccagtgg
ttgcagtc atgtcaacaggg aatgagctgc taaatcctga agatgacctc ttaccagga
agattcgaga cagcaatcgt tcaactctc tagcaacaat tcaggaacat ggttaccaca
cgatcaactt gggatttga ggagacaacc cagatgactt actcaatgcc ttgaatgagg
gtatcagtcg tgctgatgtc atcatcacat cagggggtgt atccatgggg gaaaaggact
atctcaagca ggtgctggac attgatctc atgctcagat ccatttggc agggtttta
tgaaaccagg ctgccaaca acattgcaa ctttgatat tgatggtgta agaaaaataa
tctttgact acctgggaat cctgtatcg ctgtggtcac ctgcaatctc ttgtgtgc
ctgcactgag gaaatgcag ggcacttgg atctcggcc aaccatcatc aaagcaagg
tatcatgtga tgtaaaact gatcctcgtc cagaatacca tcggtgtata ctaactggc
atcaccaaga accactacct tgggcacaga gtacaggtaa tcaatgagc agccgtctga
tgagcatgcg cagtccaat ggattgtga tgctacctc aaagacagaa cagtacgtg
agctccaca aggcgaggtg gtggatgtca tggctattg acggctatga tggtcaccag
ctgtgacaa ttaatcatc gctcgtataa tgtgtggaat tgtgagcgga taacaattc
acacaggaaa caggatccga taatgacatg cccaccgtgc ccagcacctg aactcctggg
gggaccgtca gtctctct tcccccaaa acccaaggac acctcatga tctccggac
ccctgaggtc acatgcgtg tggtagcgt gagccacgaa caccctgagg tcaagttcaa
ctggtacgtg gacggcgtg aggtgcataa tgccaagaca aagccgcggg aggagcagta
caacagcacg taccgggtg tcagcgtct caccgtctg caccaggact ggctgaatg
caaggagtac aagtgaagg tctcaacaa agccctcca gccccatc agaaaaccat
ctcaaagcc aaagggcag cccgagaacc acaggtgtac acctgccc catccggga
ggagatgacc aagaaccagg tcagcctgac ctgcctggtc aaaggctct atcccagca

catcgccgtg gagtgggaga gcaatgggca gccggagaac aactacaaga ccacgcctcc
cgtgctggac tccgacggct ccttcttct ctatagcaag ctcaccgtgg acaagagcag
gtggcagcag gggaacgtct tctcatgctc cgtgatgcat gaggctctgc acaaccacta
cacgcagaag agcctctccc tgtccccggg taaataatag gatcc

Seq. 6

ccatccagtg ctacttgtgt ttacttctaa acagtcattt tctaactgaa gctggcattc
atgtcttcat ttgggctgt ttacgtgcag ggcttcttaa aacagaagcc aactgggtga
atgtaataag tgattgaaa aaaattgaag atcttattca atctatgcat attgatgcta
ctttatatac ggaaagtgat gttcacccca gttgcaaagt aacagcaatg aagtgccttc
tcttgaggtt acaagttatt tcaattgagt ccggagatgc aagtattcat gatacagtag
aaaatctgat catcctagca aacaacagtt tgtcttctaa tgggaatgta acagaatctg
gatgcaaaga atgtgaggaa ctagaggaaa aaaatattaa agaattttg cagagttttg
tacaattgt ccaaattgtc atcaaacatt ctt ggc ggc ggc ggc ggc
gcggccgct cgaccgcggn cgcgtcgacg gcaaaggaga ctgctgaggc tgttgctgat
ganatactgg anaaggctgg gccactgtt gctgtgtctg ctgttgcaact tgatataact
gcctaccctt aaaagccaaa
ctgttgacaa ttaatcatcg gctcgtataa tgttggaat tgtgagcgga taacaatttc
acacaggaaa caggatccga taatgacatg cccaccgtgc ccagcacctg aactcctggg
gggaccgtca gtcttctct tcccccaaa acccaaggac accctcatga tctccggac
ccctgaggtc acatgcgtgg tgggtggacgt gagccacgaa gaccctgagg tcaagttcaa
ctggtacgtg gacggcgtgg aggtgcataa tgccaagaca aagccgcggg aggagcagta
caacagcacg taccgggtgg tcagcgtcct caccgtcctg caccaggact ggctgaatgg
caaggagtac aagtgaagg tctccaacaa agccctcca gccccatcg agaaaaccat
ctccaaagcc aaagggcagc cccgagaacc acagggtgac accctgcccc catcccggga
ggagatgacc aagaaccagg tcagcctgac ctgcctggc aaaggcttct atcccagcga
catcgccgtg gagtgggaga gcaatgggca gccggagaac aactacaaga ccacgcctcc
cgtgctggac tccgacggct ccttcttct ctatagcaag ctcaccgtgg acaagagcag
gtggcagcag gggaacgtct tctcatgctc cgtgatgcat gaggctctgc acaaccacta
cacgcagaag agcctctccc tgtccccggg taaataatag gatcc

Seq. 7

atgcctactt caagttctac aaagaaaaca cagctacaac tggagcattt actgctggat
ttacagatga tttgaatgg aattaataat tacaagaatc ccaaactcac caggatgctc
acatttaagt ttacatgcc caagaaggcc acagaactga aacatcttca gtgtctagaa
gaagaactca aacctctgga ggaagtgcta aatttagctc aaagcaaaaa ctttactta
agaccaggg acttaatcag caatatcaac gtaatagttc tggaactaaa gggatctgaa
acaacattca tgtgtgaata tgctgatgag acagcaacca ttgtagaatt tctgaacaga
tggtattacct tttctcaaag catcatctca aactgactt gataaggcggcggcggcggc

gcggccgcgt cgaccgcggn cgcgtcgacg gcaaaggaga ctgctgaggc tgttgctgat
ganatactgg anaaggctgg gccacttggt gctgtgtctg ctgttgcaact tgatataact
gcctacccct aaaagccaaa
ctgttgacaa ttaatcatcg gctcgtataa tgtgtggaat tgtgagcgga taacaatttc
acacaggaaa caggatccga taatgacatg cccaccgtgc ccagcacctg aactcctggg
gggaccgtca gtcttctct tcccccaaa acccaaggac accctcatga tctcccggac
ccctgaggtc acatgcgtgg tgggtggacgt gagccacgaa gaccctgagg tcaagttcaa
ctggtacgtg gacggcgtgg aggtgcataa tgccaagaca aagccgcggg aggagcagta
caacagcacg taccgggtgg tcagcgtcct caccgtcctg caccaggact ggctgaatgg
caaggagtac aagtgaagg tctccaacaa agccctccca gccccatcg agaaaaccat
ctccaaagcc aaagggcagc cccgagaacc acagggtlac accctgcccc catcccggga
ggagatgacc aagaaccagg tcagcctgac ctgcctggtc aaaggcttct atcccagcga
catcgccgtg gagtgggaga gcaatgggca gccggagaaac aactacaaga ccacgcctcc
cgtgctggac tccgacggct ccttcttct ctatagcaag ctaccgtgg acaagagcag
gtggcagcag gggaacgtct tctcatgctc cgtgatgcat gaggctctgc acaaccacta
cacgcagaag agcctctccc tgtccccggg taaataatag gatcc

Seq. 8

ccatccagt ctacttgtt ttacttctaa acagtcattt tctaactgaa gctggcattc
atgtcttcat ttgggctgt ttcagtgcag ggcttcttaa aacagaagcc aactgggtga
atgtaataag tgatttgaaa aaaattgaag atcttattca atctatgcat attgatgcta
ctttatatac ggaaagtgat gttcacccca gttgcaaagt aacagcaatg aagtgtttc
tcttgagtt acaagttatt tcacttgagt ccggagatgc aagtattcat gatacagtag
aaaatctgat catcctagca aacaacagtt tgtcttctaa tgggaatgta acagaatctg
gatgcaaaga atgtgaggaa ctgagggaaa aaaatatata agaattttg cagagtttg
tacatattgt ccaaatgttc atcaacactt ctt ggc ggc ggc ggc ggc
aatgtcccga attcccagcc tcaccacccc ttctcagtaa tgaccctggg tgggtgcagg
aggtacctac tccatactga gggtgaaatt aagggaaggc aaagtccagg cacaagagt
ggacccagc ctctcactct cagttccact catccaactg ggaccctcac cacgaatctc
atgatctgat tccgttccct gtctcctcct cccgtcacag atgtgagcca gggcactgct
cagctgtgac cctaggtgtt tctgccttgt tgacatggag agagcccttt ccctgagaa
ggcctggccc ctctctgtgc tgagcccaca gcagcaggct ggggtgtctg gttgtcagt
gtggcaccag gatggaaggg caaggcacc agggcaggcc cacagtcccg ctgtcccca
cttgaccct agctttagc tgccaacctc ccagacagcc cagcccgtg ctgagctcca
catgcatagt atcagccctc cacaccgac aaaggggaac acacccctt ggaaatggt
ctttcccc agtcccagct ggaagcatg ctgtctgttc tgctggagca gctgaacata
tacatagatg ttgccctgcc ctccccatct gcaccctgtt gagtttagt tggattgtc
tgtttatgct tggattcacc agagtgacta tgatagtga aagaaaaaaa aaaaaaaaaa

aggacgcatg tatcttgaaa tgcttgtaaa gaggtttcta acccaccctc acgaggtgtc
tctaccccc acactgggac tcgtgtggcc tgtgtggtgc caccctgctg gggcctccca
agttttgaaa ggctttcctc agcacctggg acccaacaga gaccagcttc tagcagctaa
ggaggccgtt cagctgtgac gaaggcctga agcacaggat taggactgaa gcgatgatgt
ccccttcctt acttcccctt ggggctccct gtgtcagggc acagactagg tctgtggct
ggctctggctt gcggcgcgag gatggttctc tctggtcata gccgaagtc tcatggcagt
cccaaaggag gcttacaact cctgcatcac aagaaaaagg aagccactgc cagctggggg
gatctgcagc tcccagaagc tccgtgagcc tcagccaccc ctacagactgg gttcctctcc
aagctcgccc tctggagggg cagcgcagcc tcccaccaag gggcctgcga ccacagcagg
gattgggatg aattgcctgt cctggatctg ctctagaggc ccaagctgcc tgcctgagga
aggatgactt gacaagtcag gagacactgt tcccaaagcc ttgaccagag cacctcagcc
cgctgacctt gcacaaactc catctgctgc catgagaaaa gggaagccgc ctttgcaaaa
cattgctgcc taaagaaact cagcagcctc aggcccaatt ctgccacttc tggtttgggt
acagttaaag gcaaccctga gggacttggc agtagaaatc cagggcctcc cctggggctg
gcagcttcgt gtgcagctag agctttacct gaaaggaagt ctctgggccc agaactctcc
accaagagcc tccctgccgt tcgtgagtc ccagcaattc tctaagtg aagggatctg
agaaggagaa ggaaatgtgg ggtagattg gtggtggta gagatatgcc cccctcatta
ctgccaacag ttctggctgc attcttcac gcacctcgtt tctcttctt gaagtcttg
tgccctgctc ttcagcacca tgggccttct tatacgggaag gctctgggat ctccccctg
tggggcaggc tctggggcc agcctaagat catggttag ggtgatcagt gctggcagat
aaattgaaaa ggcacgctgg cttgtgatc taaatgagga caatcccccc agggctgggc
actcctcccc tcccctcact tctccacct gcagagccag tgccttggg tgggctagat
aggatatact gtatgccggc tcttcaagc tgctgactca ctttatcaat agttccattt
aaattgactt cagtggtag actgtatcct gtttgctatt gctgttgtg ctatggggg
aggggggagg aatgtgtaag atagttaaca tgggcaaagg gagatcttg ggtgcagcac
ttaaactgcc tcgtaaccct ttcatgatt tcaaccacat ttgctagagg gagggagcag
ccacggagtt agaggccctt ggggtttctc tttccactg acaggcttc ccaggcagct
ggctagtta tcccctccc agccaggtgc aggcgttaga atatggacat ctggttgctt
tggcctgctg ccctcttca ggggtcctaa gccacaatc atgcctccct aagacctgg
catccttccc tctaagccgt tggcacctct gtgccacctc tcacactggc tccagacaca
cagcctgtgc ttttgagct gagatcactc gcttcaccct cctcatctt gttctccaag
taaagccacg aggtcggggc gagggcagag gtgatcacct gcgtgtocca totacagacc
tgcagcttca taaaacttct gatttctct cagctttgaa aagggttacc ctgggcactg
gcctagagcc tcacctcta atagacttag ccccatgagt ttgcatgtt gagcaggact
atttctggca cttgcaagtc ccatgattc ttcggttaatt ctgaggggtg ggggagggac
atgaaatcat cttagcttag ctttctgtct gtgaatgtct atatagtga ttgtgtgtt
taacaaatga ttacactga ctgttgctgt aaaagtgaat ttggaaataa agttattact
ctgatt
ctgttgacaa ttaatcatcg gctcgtataa tgtgtggaat tgtgagcgga taacaattc
acacaggaaa caggatccga taatgacatg cccaccgtgc ccagcacctg aactcctggg

ggaccgtca gtcttctct tcccccaaa acccaaggac accctcatga tctccggac
ccctgaggtc acatgctgg tggaggacgt gagccacgaa gaccctgagg tcaagttcaa
ctggtacgtg gacggcgtgg aggtgcataa tgccaagaca aagccgcggg aggagcagta
caacagcacg taccgggtgg tcagcgtcct caccgtcctg caccaggact ggctgaatgg
caaggagtac aagtgaagg tctccaacaa agccctcca gccccatcg agaaaaccat
ctcaaagcc aaagggcagc cccgagaacc acagggtgac accctgcccc catcccgga
ggagatgacc aagaaccagg tcagcctgac ctgctggtc aaaggcttct atcccagcga
catcgccgtg gagtgggaga gcaatgggca gccggagaac aactacaaga ccacgcctcc
cgtgctggac tccgacggct ccttctctct ctatagcaag ctacccgtgg acaagagcag
gtggcagcag gggaacgtct tctcatgctc cgtgatgcat gaggctctgc acaaccacta
cacgcagaag agcctctccc tgtccccggg taaataatag gatcc

Seq. 9

atgcctactt caagttctac aaagaaaaca cagctacaac tggagcattt actgctggat
ttacagatga tttgaatgg aattaataat tacaagaatc ccaaactcac caggatgctc
acatttaagt ttacatgcc caagaaggcc acagaactga aacatcttca gtgtctagaa
gaagaactca aacctctgga ggaagtgcta aatttagctc aaagcaaaaa ctttactta
agaccaggg acttaatcag caatatcaac gtaatagttc tggaactaaa gggatctgaa
acaacattca tgtgtgaata tgctgatgag acagcaacca ttgtagaatt tctgaacaga
tggattacct ttctcaaag catcatctca aactgactt gataa ggc ggc ggc ggc ggc
aatgtcccga attcccagcc tcaccacccc ttctcagtaa tgaccctggg tgggtgcagg
aggtacctac tccatactga gggtgaaatt aagggaaggc aaagtccagg cacaagagtg
ggaccccagc ctctcactct cagttccact catccaactg ggaccctcac cacgaatctc
atgatctgat tgggttccct gtctctctct cccgtcacag atgtgagcca gggcactgct
cagctgtgac cctaggtgtt tctgccttgt tgacatggag agagcccttt cccctgagaa
ggcctggccc ctctctgtgc tgagcccaca gcagcaggct ggggtgtcttg gttgtcagt
gtggcaccag gatggaaggg caaggcaccc agggcaggcc cacagtcccg ctgtcccca
cttgaccct agctttagc tgccaacctc ccagacagcc cagcccgtg ctacgtcca
catgcatagt atcagccctc cacacccgac aaaggggaac acacccctt ggaaatgggt
ctttcccc agtcccagct ggaagccatg ctgtctgtc tgctggagca gctgaacata
taccatagtg ttgccctgcc ctccccatct gcaccctgtt gattgtagt tggattgtc
tgttatgct tggattcacc agagtgacta tgatagtga aagaaaaaaa aaaaaaaa
aggacgcatg tatctgaaa tgcttgtaa gaggtttcta acccacctc acgaggtgc
tctaccccc acactgggac tcgtgtggcc tgtgtgggtc caccctgctg gggcctcca
agtttgaaa ggcttctc agcacctggg acccaacaga gaccagctc tagcagctaa
ggaggccgtt cagctgtgac gaaggcctga agcacaggat taggactgaa gcgatgatg
ccccctccct acttccccct ggggtccct gtgtcagggc acagactagg tctgtggct
ggctgtgctt gcggcgcgag gatggttctc tctggtcata gccgaagtc tcatggcagt
cccaaaggag gcttacaact cctgcatcac aagaaaaagg aagccactgc cagctggggg

gatctgcagc tcccagaagc tccgtgagcc tcagccaccc ctgagactgg gttcctctcc
aagctcgccc tctggagggg cagcgcagcc tcccaccaag ggccctgcga ccacagcagg
gattgggatg aattgcctgt cctggatctg ctctagaggc ccaagctgcc tgcctgagga
aggatgactt gacaagtcag gagacactgt tcccaaagcc ttgaccagag cacctcagcc
cgctgacctt gcacaaactc catctgctgc catgagaaaa gggaagccgc ctttgcaaaa
cattgctgcc taaagaaact cagcagcctc aggcccaatt ctgccacttc tggtttgggt
acagttaaag gcaaccctga gggacttggc agtagaaaac cagggcctcc cctggggctg
gcagcttcgt gtgcagctag agctttacct gaaaggaagt ctctgggccc agaactctcc
accaagagcc tccctgccgt tcgctgagtc ccagcaattc tctaagttg aagggatctg
agaaggagaa ggaaatgtgg ggtagatttg gtggtggta gagatatgcc cccctcatta
ctgccaacag ttctggctgc atttctcac gcacctcggg tctcttct gaagttctg
tgccctgctc ttcagcacca tgggccttct tatacgaag gctctgggat ctccccctg
tggggcaggc tcttggggcc agcctaagat catggtttag ggtgatcagt gctggcagat
aaattgaaaa ggcacgctgg cttgtgatct taaatgagga caatcccccc agggctgggc
actcctcccc tccctcact tctccacct gcagagccag tgccttggg tgggctagat
aggatatact gtatgccggc tcttcaagc tgctgactca cttatcaat agttccatt
aaattgactt cagtggtag actgtatcct gtttgctatt gctgtgtg ctatggggg
aggggggagg aatgtgtaag atagttaaca tgggcaaagg gagatcttgg ggtgcagcac
ttaaactgcc tcgtaaccct ttcatgatt tcaaccacat ttgctagagg gagggagcag
ccacggagtt agaggccctt ggggtttctc tttccactg acaggcttc ccaggcagct
ggctagtcca ttcctcccc agccaggcgc aggcgtagga atatggacat ctggttgctt
tggcctgctg cctctttca ggggtcctaa gccacaatc atgcctcct aagaccttg
catccttccc tctaagccgt tggcacctct gtgccacctc tcacactggc tccagacaca
cagcctgtgc ttttgagct gagatcactc gttcaccct cctcatctt gttctcaag
taaagccacg aggtcggggc gagggcagag gtgatcacct gcgtgtcca tctacagacc
tgcagcttca taaaacttct gatttctt cagcttgaa aagggttacc ctgggcactg
gcctagagcc tcacctcta atagacttag ccccatgagt ttgcatgtt gagcaggact
atttctggca cttgcaagtc ccatgattc ttcgtaatt ctgaggggtg ggggaggac
atgaaatcat cttagcttag ctttctgt gtgaatgtct atatagtga ttgtgtgtt
taacaaatga ttacactga ctgttctgt aaaagtgaat ttggaaataa agttattact
ctgatt
ctgttgacaa ttaatcatcg gctcgtataa tgtgtggaat tgtgagcga taacaattc
acacaggaaa caggatccga taatgacatg cccaccgtgc ccagcacctg aactcctggg
gggaccgtca gtcttctct tcccccaaa acccaaggac accctcatga tctccggac
ccctgaggtc acatgcgtgg tggtagacgt gagccacgaa gaccctgagg tcaagttcaa
ctggtacgtg gacggcgtgg aggtgcataa tgccaagaca aagccgcggg aggagcagta
caacagcacg taccgggtg tcagcgtcct caccgtcctg caccaggact ggctgaatgg
caaggagtac aagtgaagg tctccaacaa agccctcca gccccatcg agaaaacat
ctcaaagcc aaagggcagc cccgagaacc acagggtac accctgccc catccggga

ggagatgacc aagaaccagg tcagcctgac ctgcctggc aaaggcttct atcccagcga
catcgccgtg gagtgggaga gcaatgggca gccggagaac aactacaaga ccacgcctcc

cgtgctggac tccgacggct ccttcttct ctatagcaag ctcaccgtgg acaagagcag
gtggcagcag gggaacgtct tctcatgctc cgtgatgcat gaggctctgc acaaccacta
cacgcagaag agcctctccc tgtccccggg taaataatag gatcc

Seq. 10

tgtgcgcaa cacgatgaag ctcaacaaaa cgctttttat caagtcttaa atatgcctaa
cttaaatgct gatcaacgca atggttttat ccaaagcctt aaagatgac caagccaaag
tgctaacgtt ttaggtgaag ctaaaaaatt aaacgaatct caagcaccga aagctgacaa
caatttcaac aaagaacaac aaaatgcttt ctatgaaatc ttgaacatgc ctaacttgaa
cgaagaacaa cgcaatgggt tcatccaaag cttaaagat gaccaagtc aaagtgttaa
cctattgtca gaagctaaaa agttaaatga atctcaagca ccgaaagcgg ataacaaatt
caacaaagaa caacaaaatg cttctatga aatcttacct ttacctaact taaacgaaga
acaacgcaat ggtttcatcc aaagcctaaa agatgacca agccaaagcg ctaacctttt
agcagaagct aaaaagctaa atgatgcaca agcaccaaaa gctgacaaca aattcaacaa
agaacaacaa aatgctttct atgaaatttt acatttacct aacttaactg aagagcaacg
taacggcttc atccaaagcc ttaaagacga tccttcagt agcaaagaaa ttttagcaga
agctaaaaag cttaacgatg ctcaagcacc aaaagaggaa gacaacaaca aacctggtaa
agaagacggc acaaacctg gcaaagaaga cggtaac ggc ggc ggc ggc ggc
gcggccgctg cgaccgcggn cgcgtcgacg gcaaaggaga ctgctgaggc tgttctgat
ganatactgg anaaggctgg gccacttgtt gctgtgtctg ctgttgcaact tgatataact
gcctaccctt aaaagccaaa
ctgttgacaa ttaatcatg gctcgtataa tgtgtggaat tgtgagcgga taacaatttc
acacaggaaa caggatccga taatgacatg cccaccgtgc ccagcacctg aactcctggg
gggaccgtca gtcttctct tcccccaaa acccaaggac accctcatga tctcccgac
ccctgaggtc acatgcgtgg tgggtggacgt gagccacgaa gaccctgagg tcaagttcaa
ctggtacgtg gacggcgtgg aggtgcataa tgccaagaca aagccgcggg aggagcagta
caacagcacg taccgggtgg tcagcgtcct caccgtcctg caccaggact ggctgaatgg
caaggagtac aagtgaagg tctccaacaa agccctcca gccccatcg agaaaaccat
ctcaaagcc aaagggcagc cccgagaacc acagggtgac accctgcccc catcccggga
ggagatgacc aagaaccagg tcagcctgac ctgcctggc aaaggcttct atcccagcga
catcgccgtg gagtgggaga gcaatgggca gccggagaac aactacaaga ccacgcctcc
cgtgctggac tccgacggct ccttcttct ctatagcaag ctcaccgtgg acaagagcag
gtggcagcag gggaacgtct tctcatgctc cgtgatgcat gaggctctgc acaaccacta
cacgcagaag agcctctccc tgtccccggg taaataatag gatcc

Seq. 11

tgctgcgcaa cacgatgaag ctcaacaaaa cgctttttat caagtcttaa atatgcctaa
cttaaagtct gatcaacgca atggttttat ccaaagcctt aaagatgac caagccaaag

tgctaacgtt ttaggtgaag ctaaaaaatt aaacgaatct caagcaccga aagctgacaa
caattcaac aaagaacaac aaaatgcttt ctatgaaatc ttgaacatgc ctaactgaa
cgaagaacaa cgcaatggtt tcatccaaag cttaaagat gaccaagtc aaagtgctaa
cctattgtca gaagctaaaa agttaaatga atctcaagca ccgaaagcgg ataacaaatt
caacaaagaa caacaaaatg ctttctatga aatcttacat ttacctaact taaacgaaga
acaacgcaat ggtttcatcc aaagcctaaa agatgacca agccaaagcg ctaaccttt
agcagaagct aaaaagctaa atgatgcaca agcaccaaaa gctgacaaca aattcaacaa
agaacaacaa aatgctttct atgaaatctt acatttacct aacttaactg aagagcaacg
taacggcttc atccaaagcc ttaaagacga tccttcagtg agcaaagaaa ttttagcaga
agctaaaaag ctaaacgatg ctcaagcacc aaaagaggaa gacaacaaca aacctggtaa
agaagacggc acaaacctg gcaaagaaga cggtaac ggc ggc ggc ggc ggc
gttaggtca cagtgcctgc gatcacca aggtggctag aagacatgc atgtctcct
ttcctctgac atctatggac aaagcctta tcacagtcct ggagatgact ccggtgctg
ggacagaaat catcaattac cgagatggaa tggggcgagt ccttgctcaa gatgtatatg
caaaagacaa ttaccccc ttccagcat cagtaaaaga tggctatgct gccgagctg
ctgatggccc aggagatcgt tcatcattg gggaatcca agctggtgaa cagccaactc
agacagtaat gccaggacaa gtcatgcggg ttacaacagg tgctccaata ccctgcggtg
ctgatgcagt agtacaagtg gaagataccg aacttatcag ggaatcagat gatggcactg
aagaactga agtgcaatt ctggtgcaag ctggccagg ccaagatct agacccatcg
gccatgacat taaaagaggg gaatgtgtt tggccaaagg aaccacatg ggccccctag
agattggtct tctggcaact gtaggtgtca cagaggttga agttaataag tttcagtg
ttgcagtc atgcaacagg aatgagctgc taaatcctga agatgacct ttaccagga
agattcgaga cagcaatcgt tcaactctc tagcaacaat tcaggaacat ggttaccoca
cgatcaactt gggatttga ggagacaacc cagatgactt actcaatgcc ttgaatgagg
gtatcagtcg tgctgatgtc atcatcacat cagggggtgt atccatgggg gaaaaggact
atctcaagca ggtgctggac attgatctc atgctcagat ccatttggc agggtttta
tgaaaccagg ctgccaaca acatttcaa ctttgatat tgatggtgta agaaaaataa
tcttgcact acctgggaat cctgtatcgg ctgtggtcac ctgcaatctc tttgtgtgc
ctgcactgag gaaaatgcag ggcatttgg atcctcggcc aacctcatc aaagcaaggt
tatcatgtga tgtaaaact gatcctcgtc cagaatacca tcggtgtata ctaactggc
atcaccaaga accactacct tgggcacaga gtacaggtaa tcaatgagc agccgtctga
tgagcatgcg cagtccaat ggattgttga tgctacctc aaagacagaa cagtacgtgg
agctccacaa aggcgaggtg gtggatgtca tggtcattgg acggctatga tggtcaccag
ctgttgacaa ttaatcatc gctcgtataa tgtgtggaat tgtgagcgga taacaattc
acacaggaaa caggatccga taatgacatg cccaccgtgc ccagcacctg aactcctggg
gggaccgtca gtcttctct tcccccaaa acccaaggac accctcatga tctccggac

ccctgaggtc acatgcgtgg tggaggacgt gagccacgaa gaccctgagg tcaagttcaa
ctggtacgtg gacggcgtgg aggtgcataa tgccaagaca aagccgcggg aggagcagta
caacagcacg taccgggtgg tcagcgtcct caccgtcctg caccaggact ggctgaatgg
caaggagtac aagtgaagg tctccaacaa agccctcca gccccatcg agaaaacat

ctcaaagcc aaagggcagc cccgagaacc acaggtgtac accctgcccc catcccgga
ggagatgacc aagaaccagg tcagcctgac ctgcctggtc aaaggcttct atcccagca
catgccgtg gagtgggaga gcaatgggca gccggagaac aactacaaga ccacgcctcc
cgtgctggac tccgacggct cttcttctct ctatagcaag ctaccgtgg acaagagcag
gtggcagcag gggaaactct tctcatgctc cgtgatgcat gaggctctgc acaaccacta
cacgcagaag agcctctccc tgtccccggg taaataatag gatcc

Seq. 12

tgctgcgcaa cacgatgaag ctcaacaaaa cgctttttat caagtottaa atatgcctaa
cttaaagtct gatcaacgca atggttttat ccaaagcctt aaagatgac caagccaaag
tgctaacgtt ttaggtgaag ctaaaaaatt aaacgaatct caagcaccga aagctgacaa
caatttaac aaagaacaac aaaatgcttt ctatgaaatc ttgaacatgc ctaactgaa
cgaagaacaa cgcaatggtt tcatccaaag cttaaagat gaccaagtc aaagtgttaa
cctattgtca gaagctaaaa agttaaatga atctcaagca ccgaaagcgg ataacaaatt
caacaaagaa caacaaaatg cttctatga aatcttacat ttacctaact taaacgaaga
acaacgcaat ggtttcatcc aaagcctaaa agatgacca agccaaagcg ctacctttt
agcagaagct aaaaagctaa atgatgcaca agcaccaaaa gctgacaaca aattcaacaa
agaacaacaa aatgctttct atgaaatttt acatttacct aacttaactg aagagcaacg
taacggcttc atccaaagcc ttaaagacga tcttcagtg agcaaagaaa ttttagcaga
agctaaaaag ctaaacgatg ctcaagcacc aaaagaggaa gacaacaaca aacctggtaa
agaagacggc acaaacctg gcaagaaga cggtaac ggc ggc ggc ggc ggc
aatgtcccga attcccagcc tcaccacccc ttctcagtaa tgacctggt tggttgcagg
aggtacctac tccatactga gggtgaaatt aaggggaaggc aaagtccagg cacaagagt
ggaccccagc ctctactct cagttccact catccaactg ggacctcac cacgaatctc
atgatctgat tcggttccct gtctctctct cccgtcacag atgtgagcca gggcactgct
cagctgtgac cctaggtgtt tctgccttgt tgacatggag agagcccttt cccctgagaa
ggcctggccc cttcctgtgc tgagcccaca gcagcaggct ggggtgtctg gttgtcagt
gtggcaccag gatggaaggg caaggcacc agggcaggcc cacagtcccg ctgtcccca
cttgaccct agcttgtagc tgccaacctc ccagacagcc cagcccgtg ctacgtcca
catgcatagt atcagccctc cacaccgac aaaggggaac acacccctt ggaaatggt
ctttccccc agtcccagct ggaagccatg ctgtctgttc tgctggagca gctgaacata
tacatagatg ttgcctgcc ctcccatct gcacctgtt gagttgagt tggattgtc

tgtttatgct tggattcacc agagtgacta tgatagtga aagaaaaaaa aaaaaaaaaa
aggacgcatg tatctgaaa tgcttgaaa gaggtttcta acccaccctc acgaggtgtc
tctaccccc acactgggac tctgtggcc tgtgtgggc caccctgctg gggcctcca
agtttgaaa ggcttcctc agcacctggg acccaacaga gaccagcttc tagcagctaa
ggaggccgtt cagctgtgac gaaggcctga agcacaggat taggactgaa gcatgatgt
cccctccct actccctt ggggctccct gtgtcagggc acagactagg tctgtggct
ggctggctt gggcgcgag galggttctc tctggcata gcccyagtc tcatggcagt
cccaaaggag gcttacaact cctgcatcac aagaaaaagg aagccactgc cagctggggg
gatctgcagc tcccagaagc tccgtgagcc tcagccacc ctcagactgg gttcctctcc
aagctgccc tctggagggg cagcgcagcc tcccaccaag ggcctgcga ccacagcagg
gattgggatg aattgcctgt cctggatctg ctctagaggc ccaagctgcc tgctgagga
aggatgactt gacaagtcag gagacactgt tcccaaagcc ttgaccagag cacctcagcc
cgctgacctt gcacaaactc catctgtgc catgagaaaa gggaagccgc cttgcaaaa
cattgtgcc taaagaaact cagcagcctc aggcccaatt ctgccacttc tggttgggt
acagttaaag gcaaccctga gggacttggc agtagaaatc cagggcctcc cctggggctg
gcagcttctg gtgcagctag agctttacct gaaaggaagt ctctgggccc agaactctcc
accaagagcc tccctgccgt tctgtgagtc ccagcaattc tctaagttg aagggatctg
agaaggagaa ggaaatgtg ggtagattg gtggtggta gagatatgcc cccctatta
ctgccaacag ttctggctgc atttctcac gcacctcggg tctcttct gaagtcttg
tgccctgctc ttcagcacca tgggccttct tatacggaag gctctgggat ctccccctg
tggggcaggc tctggggcc agcctaagat catggttag ggtgatcagt gctggcagat
aaattgaaa ggcacgctg cttgtgatc taaatgagga caatcccccc agggctgggc
actcctccc tccctcact tctccacct gcagagccag tgccttggg tgggctagat
aggatatact gtatgccggc tcttcaagc tgctgactca cttatcaat agttccatt
aaattgactt cagtggtag actgtatct gttgtctatt gctgtgtg ctatggggg
aggggggagg aatgtgtaag atagttaaca tgggcaaagg gagatcttg ggtgcagcac
ttaaactgcc tctaacctt ttcatgatt tcaaccacat ttgctagagg gagggagcag
ccacggagt agaggccctt ggggttctc ttccactg acaggcttc ccaggcagct
ggctagtca tccctccc agccaggtgc aggcgtagga atatggacat ctggttgc
tggcctgctg cctcttca ggggtccta gccacaatc atgcctcct aagacctgg
catcctccc tctaagcgt tggcaacct gtgccacct tcacactggc tccagacaca
cagcctgtgc tttggagct gagatcact gcttaccct cctcatctt gttccaag
taaagccag aggtcggggc gagggcagag gtgatcact gcgtgtcca tctacagacc
tgcagctca taaaacttct gatttctt cagcttgaa aagggttacc ctgggcactg
gcctagagcc tcacctcta atagacttag cccatgagt ttgcatgt gagcaggact
atttctggca cttgcaagtc ccatgattc ttcggttaatt ctgaggggtg ggggaggac
atgaaatcat cttagcttag cttctgtct gtgaatgtct atatgtgta ttgtgttt
taacaaatga ttacactga ctgtgtgt aaaagtgaat ttggaataa agttattact
ctgatt

ctgttgacaa ttaatcatcg gctcgtataa tgtgtggaat tgtgagcgga taacaatttc
acacaggaaa caggatccga taatgacatg cccaccgtgc ccagcacctg aactcctggg
gggaccgtca gtcttcctct tcccccaaaa acccaaggac accctcatga tctcccggac
ccctgaggtc acatgcgtgg tggtagacgt gagccacgaa gaccctgagg tcaagttcaa
ctggtacgtg gacggcgtgg aggtgcataa tgccaagaca aagccgcggg aggagcagta
caacagcacg taccgggtgg tcagcgtcct caccgtcctg caccaggact ggctgaatgg
caaggagtac aagtgaagg tctccaacaa agccctcca gcccccatcg agaaaaccat
ctccaaagcc aaagggcagc cccgagaacc acaggtgtac accctgcccc catcccggga
ggagatgacc aagaaccagg tcagcctgac ctgcctggtc aaaggcttct atcccagcga
catgcctgtg gagtgggaga gcaatgggca gccggagAAC aactacaaga ccacgcctcc
cgtgctggac tccgacggct ccttcttct ctatagcaag ctcaccgtgg acaagagcag
gtggcagcag gggaacgtct tctcatgctc cgtgatgcat gaggctctgc acaaccacta
cacgcagaag agcctctccc tgtccccggg taaataatag gatcc